

**Tutor:** Yuehan Zhang  
[zyuehan@comp.nus.edu.sg](mailto:zyuehan@comp.nus.edu.sg)

### Theoretical:

Q1: Given a set of data points shown in figure 1, **sketch** (a schematic trend of change is enough, no need to calculate the exact numbers) a plot of how the **average square distance** of a data sample with respect to its corresponding cluster center will change as the number of clusters  $k$  increases.

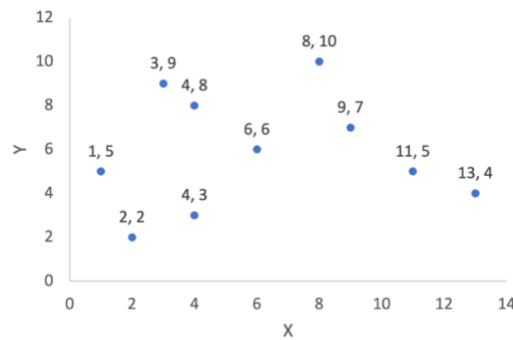


Fig 1

Q2: Consider applying SLIC superpixel segmentation to the image in Fig 2 (a). A sample result for 300 superpixels is shown in Fig 2 (b). Note how the resulting superpixels in uniform regions (e.g. those outlined in yellow) tend to be hexagonal. Why does this occur?

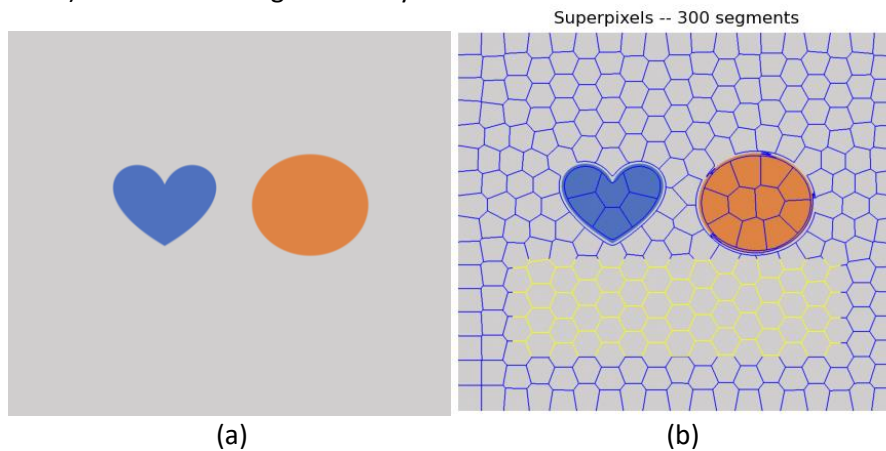


Fig 2

Q3: For Mean-Shift clustering, what happens to the number of clusters as the window size or bandwidth is increased? Explain your answer.

### Coding:

- Apply `skimage.segmentation.slic` to 100075-Bear.png and show outputs through `skimage.segmentation.mark_boundaries`. Fix sigma to 5 for all experiments.
  - Compare the results for 50, 100, 200, and 300 segments and discuss.
  - Fix the segment number to 100. Compare and discuss the results when compactness is set to 1, 10, and 100.
  - Fix segment number to 100 and compactness to 10. Compare and discuss the results when the maximum number of iterations is set to 1, 2, 3, 4 and 5.

- Apply `sklearn.cluster.KMeans` to the data in Q1. Change the initialization method to see its impact on final average square distance of all samples with respect to the clustering centers.

Past Quiz Questions:

- AY2122 quiz 2: Q1 (clustering)